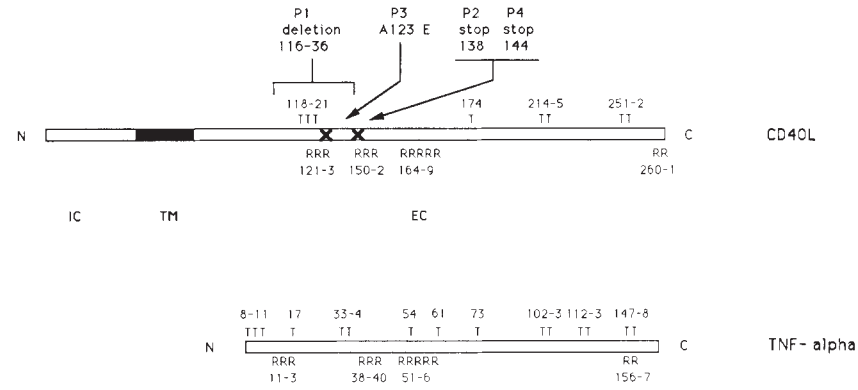


FIG. 3 Location of the CD40L mutations. CD40L compared with TNF-α with respect to sites involved in trimer formation (T) and receptor binding (R). CD40L is a type II membrane protein with extracellular (EC) C terminus and intracellular (IC) N terminus. Numbering refers to amino acids in the mature protein. Data for TNF-α taken from ref. 17. Mutations detected in the four HIGM1 patients are indicated.

in ref. 11). Moreover, isotype switch could be obtained in HIGM1 B cells after co-cultivation with Sézary's syndrome T cells[12], arguing against a primary B-cell defect. This is also supported by X-inactivation studies in obligatory carriers at HIGM1, which demonstrate a random pattern[19]. Our data correlate abnormalities of the CD40L with HIGM1 and result in the inability to bind CD40, a critical molecule for B-cell proliferation and differentiation. As a consequence, the complete signal for immunoglobulin isotype switch is not provided. Thus, HIGM1 emphasizes the central role of the CD40L/CD40 in immunoglobulin class switch. □

Received 30 December 1992; accepted 22 January 1993.

1. de Vries, J. E. *et al. Curr. Opin. Immun.* **3,** 851–858 (1991).
2. Armitage, R. J. *et al. Nature* **357,** 80–82 (1992).
3. Graf, D., Korthäuer, U., Mages, H. W., Senger, G. & Kroczek, R. A. *Eur. J. Immun.* **22,** 3191–3194 (1992).
4. Hollenbaugh, D. *et al. EMBO J.* **11,** 4313–4321 (1992).
5. Gauchat, J. F. *et al. FEBS Lett.* **315,** 259–266 (1993).
6. Gray, P. W. *et al. Nature* **312,** 721–724 (1984).
7. Wang, A. M. *et al. Science* **228,** 149–154 (1985).
8. Noelle, R. J. *et al. Proc. natn. Acad. Sci. U.S.A.* **89,** 6550–6554 (1992).
9. Mensinck, E. J. B. M. *et al. Hum. Genet.* **76,** 96–99 (1987).
10. Padayachee, M. *et al. Genomics* **14,** 551–552 (1992).
11. Notarangelo, L. D., Duse, M. & Ugazio, A. G. *Immunodef. Rev.* **3,** 101–122 (1992).
12. Mayer, L. *et al. New Engl. J. Med.* **314,** 409–413 (1986).
13. Gougeon, M. L. *et al. J. clin. Immun.* **12,** 92–100 (1992).
14. Lane, P. *et al. Eur. J. Immun.* **22,** 2573–2578 (1992).
15. Noelle, R. J., Ledbetter, J. A. & Aruffo, A. *Immun. Today* **13,** 431–433 (1992).
16. Mallett, S. & Barclay, A. N. *Immun. Today* **12,** 220–223 (1991).
17. Eck, M. J. & Sprang, S. R. *J. biol. Chem.* **264,** 17595–17605 (1989).
18. Geha, R. S. *et al. J. clin. Invest.* **64,** 385–391 (1979).
19. Hendriks, R. W. *et al. Eur. J. Immun.* **20,** 2603–2608 (1990).
20. DiSanto, J. P. *et al. Eur. J. Immun.* **23,** 320–326 (1993).
21. Winship, P. R. *Nucleic Acids Res.* **17,** 1266 (1989).

ACKNOWLEDGEMENTS. We thank F. Le Deist, A. Durandy, F. Selz and members of the clinical department of Pediatrics at Hôpital des Enfants Malades for their help in these studies, B. Bouchard and J.-P. de Villartay for review of the manuscript, and P. Lane (Basel Institute of Immunology) for the hCD40-Hμ fusion protein.

Nodal is a novel TGF-β-like gene expressed in the mouse node during gastrulation

Xunlei Zhou*, Hiroshi Sasaki†, Linda Lowe*, Brigid L. M. Hogan† & Michael R. Kuehn*‡

* Experimental Immunology Branch, National Cancer Institute, National Institutes of Health, Bethesda, Maryland 20892, USA
† Department of Cell Biology, Vanderbilt University Medical School, Nashville, Tennessee 37232, USA
‡ To whom correspondence should be addressed

DURING gastrulation, the three germ layers of the embryo are formed and organized along the anterior–posterior body axis. In the mouse, gastrulation involves the delamination of ectodermal cells through the primitive streak and their differentiation into mesoderm[1]. These processes do not occur in embryos homozygous for a retrovirally induced recessive prenatal lethal mutation, the strain 413-d insertional mutation[2,3]. Instead of giving rise to mesoderm, embryonic ectoderm in 413-d mutants overproliferates and then rapidly degenerates, although extraembryonic lineages remain viable[2]. Here we isolate a candidate for the mutated gene which encodes a new member of the transforming growth factor-β (TGF-β) superfamily[4]. Expression is first detected in primitive streak-stage embryos at about the time of mesoderm formation. It then becomes highly localized in the node at the anterior of the primitive streak. This region is analogous to chick Hensen's node and *Xenopus* dorsal lip (Spemann's organizer), which can induce secondary body axes when grafted into host embryos (reviewed in refs 5 and 6). Our findings suggest that this gene, named *nodal*, encodes a signalling molecule essential for mesoderm formation and subsequent organization of axial structures in early mouse development.

Seven overlapping cosmid clones encompassing the 413-d proviral insertion site were directly selected from a library of genomic DNA made from embryonic stem (ES) cells carrying the mutation. A map of the region flanking the provirus is shown in Fig. 1. To find the gene mutated by proviral insertion, we screened a 7.5-days-post-coitum (d.p.c.) mouse embryo complementary DNA library with a cloned genomic DNA fragment flanking the provirus. This fragment contains recognition sites for *Not*I, *Bss*HII and *Sac*II, which are characteristic of CpG islands[7,8], often found lying close to genes. From a screen of 1×10^6 recombinant phages, we isolated a single 1,800-base-pair (bp)-long cDNA clone which maps to both sides of the 413-d provirus (Fig. 1). The cDNA contains a long open reading frame (ORF) starting at the 5′ end. Although the ORF does not begin with a methionine codon, sequence analysis of cloned genomic DNA showed that the ORF continues to an in-frame Met codon 24 bp upstream. The DNA sequence of the immediate upstream genomic subclone and of the cDNA is shown in Fig. 2*a*, together with the deduced protein sequence of 354 amino acids. Examination of the protein sequence revealed extensive homology to the DVR (decapentaplegic-Vg-1-bone morphogenetic protein(BMP)-related) and activin/inhibin subgroups of the TGF-β superfamily. These secreted proteins act as signalling molecules mediating cellular interactions in many tissues during development[9]. TGF-β-like proteins are most highly conserved at their carboxy termini. This region, encompassing the mature form of the protein, is generated by proteolytic cleavage of a larger precursor. An alignment of the predicted carboxy terminal 110

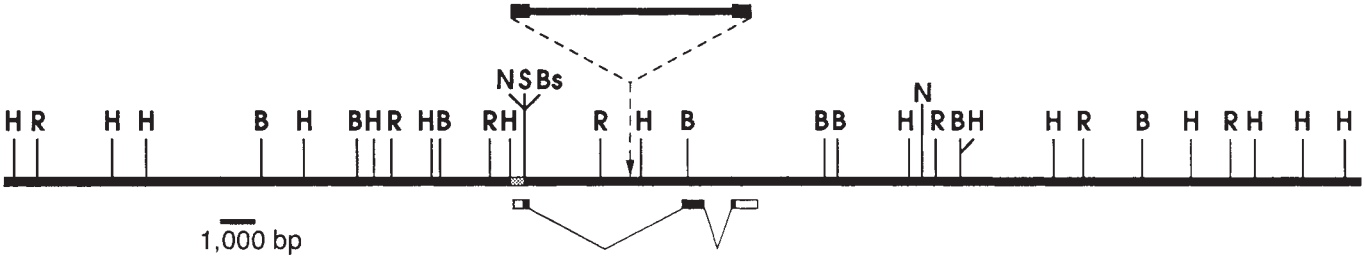


FIG. 1 Restriction map of the genomic region surrounding the site of proviral insertion (indicated by the dotted arrow). All sites for *Eco*RI (R), *Hin*dIII (H), *Bam*HI (B) and *Not*I(N), as well as the *Sac*II (S) and *Bss*HII (Bs) sites in the CpG island are shown. The genomic structure of the *nodal* cDNA is shown below. Filled and open boxes represent protein-coding regions and untranslated regions, respectively. The sequence used to probe the cDNA library is represented by the hatched region between *Hin*dIII and *Not*I sites.

METHODS. The cosmid vector pCosDN was derived by deletion of a 473-bp *Ban*II fragment from within the *neo* gene of sCos-1 (ref. 18). Genomic DNA suitable for cosmid library construction was prepared[19] from an ES cell line made from an 129/Sv embryo carrying the proviral insertion. The cosmid library was prepared essentially as described[18], using packaging extracts from Stratagene and DH10B bacteria (Gibco). The entire library, 8×10^5 colony forming units (c.f.u.), was plated onto nylon filters using ampicillin selection. To isolate clones of the proviral insertion region, we took advantage of the functional *neo* gene in the provirus[20]. One quarter of the library (2×10^5 c.f.u.) was replica-plated in kanamycin, allowing direct selection of 7 clones containing the provirus. Restriction maps of these clones were constructed using a partial digestion mapping procedure[18]. Fragments were subcloned into pBluescript SK (Stratagene) using standard procedures. The restriction map was confirmed by hybridization of several unique sequence flanking probes to Southern blots of 129/Sv genomic DNA (data not shown).

a

```
   1  AAGCTTCACTCGGAGCAAGCCTTAGCCCGCTGTCTCAGCAGGGAGACTTCCCGAGGTAGAGGGGCAAGGTGCGGGGCGGTTTAGACTCAGAGTCTGTATG

 101  CACCCCTAACTCCCCCCCCCCCCCCGCCACAATTTCTCTGTAGTCTTTCTCAGCACATCACACCTCCCCTCAGCAGGGGCTCCCCTGCCCTGCCCCTCC

 201  AGGGTGGTTATAAGTTCTTAACCTATAGGTTATAGGCCTCTCCGGAGGGAGGGAGGGAAAGGGGCGGGGCGGCGGCGGCTCAGATATAAGGGACCCTGGT

 301  GTTGTTTGCGGGTCCAAACAGCCCACC ATG AGT GCC CAC AGC CTC CGC ATC CTT CTT CTT CAA GCC TGT TGG GCT CTA CTC
   1                             Met Ser Ala His Ser Leu Arg Ile Leu Leu Leu Gln Ala Cys Trp Ala Leu Leu

 382  CAC CCG CGC GCC CCG ACC GCG GCC GCT TTG CCT CTG TGG ACA CGG GGG CAG CCC TCG TCA CCG TCC CCT CTG GCG
  19  His Pro Arg Ala Pro Thr Ala Ala Ala Leu Pro Leu Trp Thr Arg Gly Gln Pro Ser Ser Pro Ser Pro Leu Ala
                                                                                    ↓
 457  TAC ATG TTG AGC CTC TAC CGA GAC CCG CTG CCT CGG GCG GAC ATC ATC CGC AGC CTC CAG GCG CAA GAT GTG GAC
  44  Tyr Met Leu Ser Leu Tyr Arg Asp Pro Leu Pro Arg Ala Asp Ile Ile Arg Ser Leu Gln Ala Gln Asp Val Asp

 532  GTG ACC GGA CAG AAC TGG ACT TTC ACG TTT GAC TTC TCC TTT TTG AGC CAA GAA GAG GAT CTG GTA TGG GCG GAC
  69  Val Thr Gly Gln Asn Trp Thr Phe Thr Phe Asp Phe Ser Phe Leu Ser Gln Glu Glu Asp Leu Val Trp Ala Asp

 607  GTC CGG TTG CAG CTG CCG GGC CCC ATG GAC ATA CCC ACT GAG GGC CCA CTC ACC ATT GAC ATT TTC CAC CAG GCC
  94  Val Arg Leu Gln Leu Pro Gly Pro Met Asp Ile Pro Thr Glu Gly Pro Leu Thr Ile Asp Ile Phe His Gln Ala

 682  AAG GGG GAT CCA GAG CGG GAC CCC GCT GAC TGC CTG GAG CGC ATT TGG ATG GAG ACG TTC ACC GTC ATT CCT TCT
 119  Lys Gly Asp Pro Glu Arg Asp Pro Ala Asp Cys Leu Glu Arg Ile Trp Met Glu Thr Phe Thr Val Ile Pro Ser

 757  CAG GTC ACG TTT GCC TCA GGC AGC ACA GTC CTG GAG GTG ACC AAG CCA CTC TCC AAG TGG CTA AAG GAC CCC AGG
 144  Gln Val Thr Phe Ala Ser Gly Ser Thr Val Leu Glu Val Thr Lys Pro Leu Ser Lys Trp Leu Lys Asp Pro Arg

 832  GCA CTG GAA AAG CAG GTG TCC AGT CGA GCA GAA AAG TGT TGG CAT CAG CCC TAC ACC CCA CCT GTA CCT GTC GCC
 169  Ala Leu Glu Lys Gln Val Ser Ser Arg Ala Glu Lys Cys Trp His Gln Pro Tyr Thr Pro Pro Val Pro Val Ala

 907  AGC ACC AAT GTG CTC ATG CTC TAC TCC AAC CGG CCT CAG GAG CAG AGG CAG CTA GGG GGC GCC ACT TTG CTT TGG
 194  Ser Thr Asn Val Leu Met Leu Tyr Ser Asn Arg Pro Gln Glu Gln Arg Gln Leu Gly Gly Ala Thr Leu Leu Trp

 982  GAA GCT GAG AGC TCC TGG CGG GCC CAG GAG GGA CAG CTG TCT GTA GAG AGG GGC GGA TGG GGC AGA AGG CAA CGC
 219  Glu Ala Glu Ser Ser Trp Arg Ala Gln Glu Gly Gln Leu Ser Val Glu Arg Gly Gly Trp Gly Arg Arg Gln Arg

1057  CGA CAT CAT TTG CCA GAC AGA AGC CAA CTG TGT AGG AGG GTC AAG TTC CAG GTG GAC TTC AAC CTG ATT GGC TGG
 244  Arg His His Leu Pro Asp Arg Ser Gln Leu Cys Arg Arg Val Lys Phe Gln Val Asp Phe Asn Leu Ile Gly Trp

1132  GGC TCC TGG ATC ATC TAC CCC AAG CAG TAC AAT GCC TAT CGC TGT GAG GGC GAG TGT CCT AAC CCT GTG GGG GAG
 269  Gly Ser Trp Ile Ile Tyr Pro Lys Gln Tyr Asn Ala Tyr Arg Cys Glu Gly Glu Cys Pro Asn Pro Val Gly Glu
                                    ↓
1207  GAG TTT CAT CCT ACC AAC CAT GCC TAC ATC CAG AGC CTG CTG AAA CGA TAC CAA CCC CAC CGG GTT CCT TCC ACG
 294  Glu Phe His Pro Thr Asn His Ala Tyr Ile Gln Ser Leu Leu Lys Arg Tyr Gln Pro His Arg Val Pro Ser Thr

1282  TGC TGT GCC CCC GTG AAG ACC AAG CCA CTG AGC ATG CTT TAT GTG GAC AAT GGC AGG GTC CTC CTG GAA CAC CAC
 319  Cys Cys Ala Pro Val Lys Thr Lys Pro Leu Ser Met Leu Tyr Val Asp Asn Gly Arg Val Leu Leu Glu His His

1357  AAG GAC ATG ATT GTG GAG GAG TGT GGG TGC CTC TGA CAGAGCCAGGGGGAGTGCTGAAATTGGCTTGCATTCCACAATGCTGATGAAC
 344  Lys Asp Met Ile Val Glu Glu Cys Gly Cys Leu ***

1445  TCCAAGGAGACTCCATTGTGTCTATCCAGGGAGCAGAAACGTTAGAAGAGTTCTGCCTGCTGGAGCTAAAGAGAAAAGCCCCGCCCCCTGTGCATACAGT

1545  GCTCTTAGACCTGCCAAGCCAGAGAGAGGCTACCGTGGCATGGCAGGATGGGGAAGCCTTGCAGGGGCTGGCTCGCTGGGCTCCCTGGAAATAGGGTTTA

1645  TGAACTGCTTGAAATTGTGTGCAAAGGCTGGGGTGTATATATATATATGTATATATATATATGTATTTATTTTTGATGCCATCTGTAACTTTCTTGTA

1745  TTCTCAAAAGTGGTCTGTGACCTGCTGTCCCTCCCTCAAGATTAGTATATATTTTATTAGATTATAAACGAGCCATTTGGTTCTCCCTGCCTCAAGCTGT

1845  GGTAGGGAAGACCCACAACCTTCTGGCTGGCTGGCAGTGACATCCTGGCCTTGGTCAGGGGCTCTCTGATCTCTAATGACTTGCCTAAAAAAGCCACTGT

1945  CCAGTTCTCCAGGGCCAGTTGGTGCCTTTGACCAGAGAGGTGGGCACTTGTCCAAGAGGGGACTGGCCATGGTGGACTTTAGAAGCCAGAGTCCTGAGAT

2045  GTATGCTTGGCAGACACAACCCAAGTCTATTAAAAGTCTGTGACAATTCAAAAAAAAAAAAAAAAAAAAAAAAAAA
```

amino acids of this new gene, which we have named *nodal* (see later), with those of several TGF-β superfamily members is shown in Fig. 2*b*. Shared structural features include a cluster of basic amino acids at the start of the conserved region, believed to be important for proteolytic cleavage, and seven highly conserved cysteine residues. Overall, *nodal* protein is from 34–39% identical in the conserved region to other superfamily members, and about 25% identical to TGF-β itself. The deduced evolutionary relationships among superfamily members are shown in Fig. 2*c*. The first 20 to 30 amino acids of *nodal* have the characteristics of a signal sequence, suggesting that *nodal* encodes a secreted protein.

To determine the *in vivo* expression pattern of *nodal*, both reverse-transcriptase PCR (RT-PCR)[10] and whole-mount *in situ* hybridization were carried out on wild-type embryos at about the time of gastrulation. In the mouse, gastrulation starts at ~6.5 d.p.c. with the formation of the primitive streak near the junction of the embryonic and extraembryonic ectoderm at the posterior of the embryo. As development proceeds, the primitive streak elongates owing to growth of the egg cylinder and movement of cells from the embryonic ectoderm into the posterior of the streak. As a result, the anterior of the streak comes to lie at the distal tip of the egg cylinder, where a discrete structure known as the node becomes visible at approximately 7.5 d.p.c. Cells leaving the node give rise to definitive endoderm, notochord and paraxial mesoderm[11–13]. Using RT-PCR, *nodal* RNA is detectable in total RNA prepared from pre-streak and very early streak embryos, although the amount of amplified product is far below that found for later stage embryos (Fig. 3). The first detectable expression of *nodal* by whole-mount *in situ* hybridization coincides with the appearance of the node. The hybridization is highly localized and can be seen as a ring of staining around the node in Fig. 4. This pattern persists through the neural plate/head-fold stage, but no signal is detected by 8.5 d.p.c., coinciding with the disappearance of the node as a distinct structure. It is because of this localized expression in the node that we have proposed the name *nodal* for this gene.

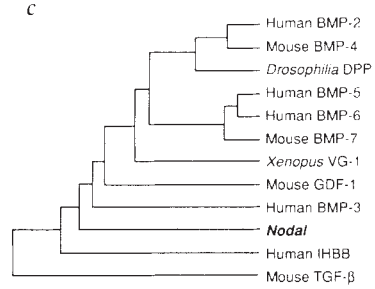
Intense effort is currently focused on the molecular mechanisms underlying mesoderm induction and gastrulation in vertebrates. In *Xenopus* the earliest event involves the induction of cells in the equatorial zone to form mesoderm by factors secreted by vegetal pole cells (reviewed in ref. 14). Members of the TGF-β, fibroblast growth factor and *wnt* families of polypeptide signalling molecules have been implicated as mesoderm-inducing factors[15]. Other extracellular factors, such as *noggin*[16], may also be involved in mesoderm formation and axial development. One problem encountered in *Xenopus* when determining the relative importance of each of these factors in mesoderm formation is the difficulty in removing them from the developing embryo and assaying the consequences. Although the mouse has limitations for experimental embryology, it does have the advantage of mutational analysis. Our study of a retroviral insertional mutant has led to the identification of the gene *nodal*, which is required for early embryonic development. Several lines of evidence suggest that *nodal* encodes an essential component of the process of mesoderm formation in the mouse. Embryos homozygous for the retroviral insertion within the *nodal* locus do not form embryonic mesoderm[2]; the *nodal* sequence predicts a potentially secreted protein related to TGF-β; and *nodal* expression is detected (by RT-PCR) in embryos at the time of primitive streak formation. It has also been shown that ES cells homozygous for the insertional mutation

b

```
Mouse nodal      RHHLP  DRS QLCRRVKFQVDFN LIGWGSWIIYPKQYNAYRCEGECPNP  VGEEFHPT
Human BMP-2      HKQRK  RLK SS_KRHPLY___S DV__ND__VA_PG_H_FY_H_E__F_  LADHLNS_
Human BMP-3      RRKQW  IEP RN_ARRYLK___A DI__SE___S_KSFD_YY_S_A_QF_  MPKSLKPS
Mouse BMP-4      PQRSR  KKN KN_RRHSLY___S DV__ND__VA_PG_Q_FY_H_D__F_  LADHLNS_
Human BMP-5      DYNTS  EQK QA_KKHELY_S_R DL__QD___A_EG_A_FY_D_E_SF_  LNAHMNA_
Human BMP-6      DYNSS  ELK TA_RKHELY_S_Q DL__QD___A_KG_A_NY_D_E_SF_  LNAHMNA_
Mouse BMP-7      ENSSS  DQR QA_KKHELY_S_R DL__QD___A_EG_A_YY_E_E_AF_  LNSYMNA_
Drosophila DPP   PTRRK  NHD DT_RRHSLY___S DV__DD__VA_LG_D_YY_H_K__F_  LADHFNS_
Xenopus VG-1     SKLPF  TAS NI_KKRHLY_E_K DV__QN_V_A_QG_M_NY_Y_E__Y_  LTEILNGS
Mouse GDF-1      PRVEV  GPV GT_RTRRLH_S_R EV__HR_V_A_RGFL_NF_Q_T_AL_ETLRGPGGPP
Human IHBB       RGLEC  DGRTNL_CRQQFFI__R LI__ND___A_TG_YGNY_E_S__AY  LAGVPGSA
Mouse TGF-β      LDTNYCFSSTEKN_CVRQLYI__RKDL__ K__HE_KG_H_NF_L_P__Y  IWS LDT
Consensus                   C    LYVDF  D GW DWIIAP GY A YC G CPFP  L    N T

Mouse nodal        NHAYIQSLLKRYQPHRVP STCCAPVKTKPLSMLYVDN GRVLLEHHKDMIVEECGCL
Human BMP-2        ___IV_T_VNSVN SKI_ KA__V_TELSAI_M__L_ENEK_V_KNYQD_V_EG___R
Human BMP-3        ___TI_SIVRAVGVVPGIPEP__V_EKMSSL_I_FF_ENKN_V_KVYPN_T_ES_A_R
Mouse BMP-4        ___IV_T_VNSVN SSI_ KA__V_TELSAI_M__L_EYDK_V_KNYQE_V_EG___R
Human BMP-5        ___IV_T_VHLMF_DHV_ KP__A_TKLNAI_V__F_DSSN_I_KKYRN_V_RS___
Human BMP-6        ___IV_T_VHLMN_EYV_ KP__A_TKLNAI_V__F_DNSN_I_KKYRN_V_RA___H
Mouse BMP-7        ___IV_T_VHFIN_DTV_ KP__A_TQLNAI_V__F_DSSN_I_KKYRN_V_RA___H
Drosophila DPP     ___VV_T_VNNMN_GKV_ KA__V_TQLDSVAM__LNDQST_V_KNYQE_T_VG___R
Xenopus VG-1       ___IL_T_VHSIE_EDI_ LP__V_TKMSPI_M_FY_NNDN_V_RHYEN_A_DE___R
Mouse GDF-1    AL___VLRA_MHAAA_TPGAGSP__V_ERLSPI_V_FF_NSDN_V_RHYED_V_DE___R
Human IHBB     SSF_TAVVNQYRMRGLNPGTVNS__I_TKLSTM_M__F_DEYNIVKRDVPN_I_EE___A
Mouse TGF-β      QYSKVLA_YNQHN_GASA SP__V_QALEPLPIV_YVG RKPKVEQLSN_I_RS_K_S
Consensus        NHA VQTLV    P   P  PCCVPT L  IS LY D   NVVLK Y NM V  CGCR
```

c



◀ FIG. 2 Sequence analysis of *nodal*. *a*, Nucleotide and deduced amino-acid sequences of the cDNA and upstream genomic subclone. The nucleotide sequence begins at the *Hin*dIII site in the genomic subclone (the left end of the hatched region in Fig. 1). The cDNA begins at the underlined C at position 352 and the overlap of genomic and cDNA sequences continues to position 408. Exon/intron junctions, determined by sequencing appropriate genomic subclones, are marked by arrows. *b*, Sequence alignment of the final 110 amino acids (single-letter code) of the predicted protein sequence of the *nodal* gene with the carboxy terminal portions of several TGF-β superfamily members[4] generated using GeneWorks (Intelligenetics). DPP, decapentaplegic; IHBB, inhibin B β-chain. The 7 invariant cysteine residues are shown in bold. These and other amino acids that do not vary from the *nodal* sequence are indicated by dashes. The consensus sequence is shown below. *c*, Tree of alignment generated by GeneWorks, using the unweighted pair group methods with arithmetic mean, showing the calculated evolutionary relationships of the sequences aligned in *b*.

METHODS. The 7.5-d.p.c. mouse embryo (CD1 strain) λZap cDNA library (gift from D. Weng and J. Gearhart) was plated at a density of 2×10^5 plaque-forming units per 22-cm^2 plate (Nunc) using 2 plates. Duplicate plaque lifts were made with GeneScreen Plus (NEN) nylon filters. Filters were hybridized with the CpG island probe (see Fig. 1) according to the manufacturer's instructions. The cDNA insert was excised from the λZap phage as a pBluescript SK subclone. Dideoxy sequencing was done on double-stranded DNA using ^{35}S-labelled dATP (NEN) and Sequenase (USB). Initial sequencing used T3 and T7 promoter primers. Subsequent sequencing primers were designed from the sequence data. This strategy was continued until the entire cDNA subclone was sequenced.

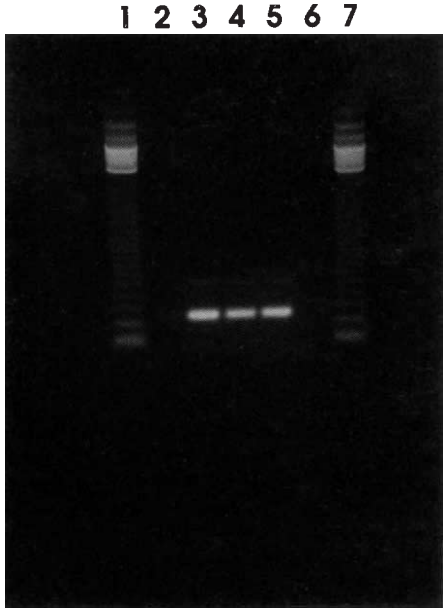


FIG. 3 Embryonic expression of *nodal* RNA. RT-PCR analysis of total RNA prepared from embryos at ~6.25 d.p.c. (pre- and very early streak stage) (lane 2); at 6.75 d.p.c. (early streak stage) (lane 3); 7.25 d.p.c. (late streak stage, but before appearance of node) (lane 4); 7.5 d.p.c. (head-fold stage with node visible) (lane 6, control (water in place of RNA); lanes 1 and 7, 123-bp ladder (Gibco). The band visible in all lanes (except control) results from amplification of a 290-bp fragment using primers lying in different exons.

METHODS. Timed pregnancies were obtained as described for Fig. 4. Reverse transcription was done on 1 μg total RNA using specific priming. The primer used, C6, is the reverse complement of nucleotides 1,721 to 1,739 in the sequence in Fig. 1. Three rounds of fully nested PCR were then performed on the cDNA samples. The first round, using NH-2 (361 to 384) as forward primer and C6 as reverse primer, was for 20 cycles, with each cycle consisting of 94 °C for 1 min, 50 °C for 1 min, and 72 °C for 1 min 15 s. The second round, using P1112 (1,112 to 1,131) as forward primer and P1447 (reverse complement of 1,427 to 1,447) as reverse primer, was also for 20 cycles, with each cycle consisting of 94 °C for 1 min, 55 °C for 1 min, and 72 °C for 1 min. The third round, using P1133 (1,133 to 1,153) as forward primer and RC-9 (reverse complement of 1,406 to 1,423) as reverse primer, was also for 20 cycles with each cycle consisting of 94 °C for 1 min, 50 °C for 1 min, and 72 °C for 1 min.

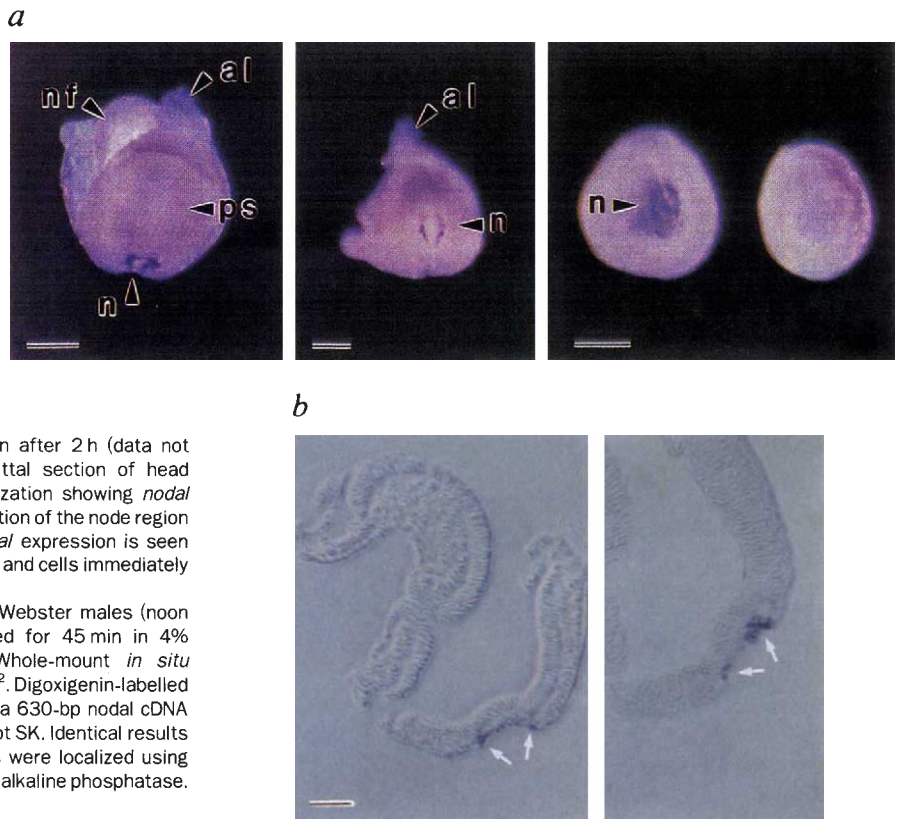


FIG. 4 Localization of *nodal* RNA. *a*, Left and middle panels, whole-mount *in situ* hybridization of two head-fold-stage embryos (~7.5 d.p.c.; no visible somites), viewed from the ventral surface with posterior end up, showing expression in the node (n). The neural folds (nf) at the anterior of the embryo, and the primitive streak (ps) and allantois (al) at the posterior, show no hybridization signal. Right panel, two later primitive streak-stage embryos (~7 d.p.c.) from the same litter. The embryo on the left shows expression associated with the node. The less developed embryo at right shows a very low level of expression and no discrete node. In all cases the colouring reaction was carried out for 15 h. No specific signal was seen with sense-strand probe. Experiments performed under essentially the same conditions revealed expression of *goosecoid*[21] RNA (from E. M. DeRobertis) in embryos before node formation after 2 h (data not shown). Scale bar, 0.2 mm. *b*, Left panel, parasagittal section of head fold-stage embryo after whole-mount *in situ* hybridization showing *nodal* expression (arrows). Right panel, high-power magnification of the node region of a more sagittal section of the same embryo. *Nodal* expression is seen in a small group of external cells (possibley endoderm) and cells immediately subjacent to them. Scale bar, 50 μm.

METHODS. ICR female mice were mated with Swiss Webster males (noon on the day of plug is 0.5 d.p.c.) and embryos fixed for 45 min in 4% paraformaldehyde in phosphate-buffered saline. Whole-mount *in situ* hybridization was carried out essentially as described[22]. Digoxigenin-labelled antisense and sense strand RNA was prepared from a 630-bp nodal cDNA fragment (nucleotides 571–1,220) cloned in pBluescript SK. Identical results were obtained with a full-length cDNA probe. Hybrids were localized using sheep anti-digoxigenin Fab fragments, conjugated with alkaline phosphatase.

differentiate normally into mesoderm in injection chimaeras made with wild-type embryos[3]. All these findings are consistent with the normal *nodal* gene product encoding a secreted signalling molecule that mediates mesoderm formation. A key observation reported here is that *nodal* gene expression becomes highly localized in the node of the mouse embryo. Evidence suggests that this region plays a crucial role in the organization of mesodermal lineages along the anterior–posterior body axis. It may contain a resident population of stem cells influenced by locally produced extracellular morphogens[13,17]. Future studies on the molecular mechanisms of *nodal* function are therefore likely to improve our understanding of both mesoderm induction and specification as well as axis formation in the mammalian embryo. □

Received 19 October; accepted 4 December 1992.

1. Bellairs, R. *Anat. Embryol.* **174,** 1–14 (1986).
2. Iannaccone, P. M., Zhou, X., Khokha, M., Boucher, D. & Kuehn, M. R. *Dev. Dynamics* **194,** 198–208 (1992).
3. Conlon, F. L., Barth, K. S. & Robertson, E. J. *Development* **111,** 969–981 (1991).
4. Massagué, J. *Rev. Cell Biol.* **6,** 597–641 (1990).
5. Nakamura, O. & Toivonen, S. *Organizer—A Milestone of a Half Century from Spemann* (Elsevier/North Holland Biomedical, Amsterdam, Oxford and New York, 1978).

6. Slack, J. M. W. *From Egg to Embryo* 2nd edn (Cambridge University Press, Cambridge, 1991).
7. Gardiner-Garden, M. & Frommer, M. *J. molec. Biol.* **196,** 261–282 (1987).
8. Bird, A. P. *Trends Genet.* **3,** 342–347 (1987).
9. Lyons, K. M., Jones, C. M. & Hogan, B. L. M. *Trends Genet.* **7,** 408–412 (1991).
10. Kawasaki, E. S. in *PCR Protocols: A Guide to Methods and Applications* (eds Innis, M. A., Gelfand, D. H., Sninsky, J. J. & White, T. J.) 21–27 (Academic, San Diego, 1990).
11. Tam, P. P. L. & Beddington, R. S. P. in *Postimplantation Development in the Mouse* (eds Chadwick, D. J. & Marsh, J.) 27–60 (Wiley, Chichester, 1992).
12. Lawson, K. A., Meneses, J. J. & Pedersen, R. A. *Development* **113,** 891–911 (1991).
13. Lawson, K. A. & Pedersen, R. A. in *Postimplantation Development in the Mouse* (eds Chadwick, D. J. & Marsh, J.) 3–20 (Wiley, Chichester, 1992).
14. Smith, J. C. *Development* **105,** 665–677 (1989).
15. Jessell, T. M. & Melton, D. A. *Cell* **68,** 257–270 (1992).
16. Smith, W. C. & Harland, R. M. *Cell* **70,** 829–840 (1992).
17. Hogan, B. L. M., Thaller, C. & Eichele, G. *Nature* **359,** 237–241 (1992).
18. Evans, G. A., Lewis, K. & Rothenberg, B. E. *Gene* **79,** 9–20 (1989).
19. Herrmann, B. G. & Frischauf, A.-M. in *Guide to Molecular Cloning Techniques* (eds Berger, S. L. & Kimmel, A. R.) 180–183 (Academic, San Diego, 1987).
20. Seliger, B., Kollek, R., Stocking, C., Franz, T. & Ostertag, W. *Molec. cell. Biol.* **6,** 286–293 (1986).
21. Blum, M. *et al. Cell* **69,** 1097–1106 (1992).
22. Conlon, R. A. & Rossant, J. *Development* **116,** 357–368 (1992).

ACKNOWLEDGEMENTS. This work was supported by grants from the NICHD to M.R.K. and B.L.M.H. H.S. is a recipient of a fellowship from the Human Frontier Science program. We thank Drs Chris Wright (Vanderbilt) and Dinah Singer (NIH) for helpful comments on the manuscript and C. Michael Jones for mouse embryo RNA.

Secreted *noggin* protein mimics the Spemann organizer in dorsalizing *Xenopus* mesoderm

William C. Smith, Anne K. Knecht, Mike Wu & Richard M. Harland*

Department of Molecular and Cell Biology, Division of Biochemistry and Molecular Biology, 401 Barker Hall, University of California, Berkeley, California 94720, USA

A DORSALIZING signal acts during gastrulation to change the specification of lateral mesodermal tissues from ventral (blood, mesenchyme) to more dorsal fates (muscle, heart, pronephros)[1–3]. This signal, from Spemann's organizer, cannot be mimicked by the mesoderm inducers activin and fibroblast growth factor[2]. The gene *noggin* is expressed in the organizer[4], and could be the dorsalizing signal. Here we show that soluble *noggin* protein added to ventral marginal zones during gastrulation induces muscle, but that activin does not. Dorsal pattern can be partially rescued in ventralized embryos by injection of a plasmid that expresses *noggin* during gastrulation. The results suggest that the *noggin* product may be the dorsalizing signal from the organizer.

Previously we examined the effects of the expression of *noggin* protein (noggin) from injected messenger RNA and concluded that noggin can act early through a vegetal dorsalizing centre to induce a Spemann organizer[4]. To study late inductive effects mediated by organizer-derived noggin, it was necessary to control the time of exposure of embryonic tissues to noggin. To prepare active, soluble noggin protein, *Xenopus* oocytes were injected with *noggin* mRNA, and secreted products collected. After incubation with ^{35}S-methionine, fluorographs of SDS-polyacrylamide gels (Fig. 1) show a major secreted protein product only after injection of *noggin* mRNA. Without a reducing agent, secreted noggin protein had an M_r of ~64,000 (64K) (Fig. 1). Reduction of disulphide bonds with β-mercaptoethanol reduced the M_r by about half to 33K, indicating that noggin is secreted as a dimer (Fig. 1). Treatment of this sample with *N*-glycanase to remove *N*-linked carbohydrate chains reduced the M_r to ~28K (Fig. 1). A minor species of intermediate M_r (30K) was also present in *N*-glycanase treated samples, suggesting that there were other modifications to the peptide. These results agree closely with the reading frame of cloned noggin cDNA[4], which predicts a secreted, glycosylated peptide of M_r ~26K, including the signal peptide.

One potential target of noggin, as a protein released from the organizer, is ventrally specified mesoderm in the marginal zone of *Xenopus* embryos. Explanted blastula and gastrula stage ventral marginal zones (VMZ) were treated with diluted noggin-conditioned oocyte medium, cultured to the late neurula stage, then assayed both for changes in morphology and for molecular markers of dorsal development. Other explants were incubated in buffer alone (see Fig. 2), with diluted medium from non-injected oocytes, or with activin. RNA isolated from the marginal zones at the end of the incubation was analysed for muscle-specific actin expression[5,6]. Without added inducers, only dorsal marginal zones (DMZ) from the gastrula stage expressed muscle actin (Fig. 2*B*). Noggin-conditioned oocyte medium induced expression of a large amount of muscle actin RNA when added to gastrula VMZs. Neither medium from non-injected oocytes, nor purified activin had this inductive activity at the gastrula stage. In contrast, both noggin and activin induced muscle actin expression in VMZs when added at the blastula stage.

Ventral marginal zones treated with noggin became elongated, which is a characteristic of dorsal development (Fig. 2*C*), although the elongation was not as extensive as DMZs. Activin induced elongation only when added at the blastula stage (not shown). Explanted gastrula stage marginal zones were immunostained with an anti-notochord antibody[7] (Fig. 2*C*). All DMZ explants stained for notochord, whereas VMZs incubated with medium from non-injected oocytes did not. Only one of nine activin-treated VMZs stained for notochord. Likewise, only one of nine VMZs incubated with noggin stained, despite their elongated appearance.

These results show that noggin is soluble, that it induces muscle in gastrula ventral mesoderm, but at this stage it induces notochord weakly or not at all. Thus noggin resembles the

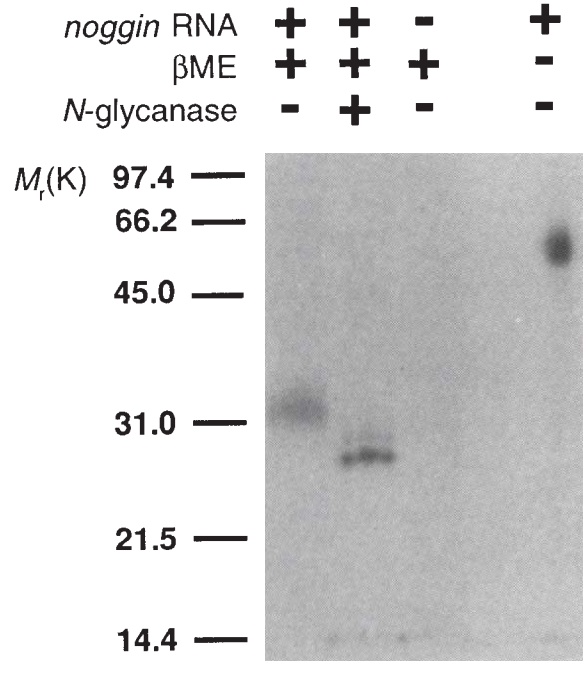


FIG. 1 Translation of noggin mRNA in *Xenopus* oocytes. Fluorographed SDS-polyacrylamide gel of medium from noggin RNA-injected *Xenopus* oocytes. Also shown is control medium from non-injected oocytes. (Samples were treated with β-mercaptoethanol (βME) or *N*-glycanase as indicated.) METHODS. Manually dissected oocytes were injected with 50 ng *noggin* Δ5′ RNA. Oocytes were cultured for 3 d at 19 °C in 20 μl per oocyte of Barth's solution (88 mM NaCl, 1 mM KCl, 0.82 mM $MgSO_4$, 2.4 mM $NaHCO_3$, 0.33 mM $Ca(NO_3)_2$, 0.41 mM $CaCl_2$ and 10 mM HEPES, pH 7.6) supplemented with 0.5 mg ml^{-1} BSA and 0.1 mCi ml^{-1} ^{35}S-methionine. Samples (1 μl) of medium were analysed by fluorography after electrophoresis on 12.5% polyacrylamide–SDS gels.

* To whom correspondence should be addressed.